The results obtained so far suggest abolishing the distinction between a context of discovery and a context of justification, norms and facts, observational terms and theoretical terms. None of these distinctions plays a role in scientific practice. Attempts to enforce them would have disastrous consequences. Popper's 'critical' rationalism fails for the same reasons.

Let us now use the material of the preceding sections to throw light on the following features of contemporary empiricism: (1) the distinctions between a context of discovery and a context of justification – norms and facts, observational terms and theoretical terms; (2) Popper's 'critical' rationalism; (3) the problem of incommensurability. The last problem will lead us back to the problem of rationality and order vs anarchism, which is the main topic of this essay.

One of the objections which may be raised against my attempt to draw methodological conclusions from historical examples is that it confounds two contexts which are essentially distinct, viz. a context of discovery, and a context of justification. *Discovery* may be irrational and need not follow any recognized method. *Justification*, on the other hand, or – to use the Holy Word of a different school – *criticism*, starts only *after* the discoveries have been made, and it proceeds in an orderly way. 'It is one thing,' writes Herbert Feigl, 'to retrace the historical origins, the psychological genesis and development, the socio-political-economic conditions for the acceptance or rejection of scientific theories; and it is quite another thing to provide a logical reconstruction of the conceptual structure and of the testing of scientific theories.'[1] These are indeed two different *things*, especially as they are done by two different *disciplines* (history of science, philosophy of science), which are quite jealous of their independence. But the question is not what distinctions a fertile mind can dream up when confronted with a complex process, or how some homogeneous material may be subdivided; the question is to what extent the distinction drawn reflects a real difference, and whether science can advance without a strong interaction between the separated

1. 'The Orthodox View of Theories', in Radner-Winokur (eds), *Analyses of Theories and Methods of Physics and Psychology*, Minneapolis, 1970, p. 4.

domains. (A river may be subdivided by national boundaries but this does not make it a discontinuous entity.) Now there is, of course, a very noticeable difference between the rules of testing as 'reconstructed' by philosophers of science and the procedures which scientists use in actual research. This difference is apparent to the most superficial examination. On the other hand, a most superficial examination also shows that a determined application of the methods of criticism and proof which are said to belong to the context of justification would wipe out science as we know it – and would never have permitted it to arise.[2] Conversely, the fact that science exists proves that these methods were frequently overruled. They were overruled by procedures which belong to the context of discovery. Thus the attempt 'to retrace the historical origins, the psychological genesis and development, the socio-political-economic conditions for the acceptance or rejection of scientific theories', far from being irrelevant for the standards of test, actually leads to a criticism of these standards – *provided* the two domains, historical research and discussion of test procedures, are not kept apart by fiat.

In another paper Feigl repeats his arguments and adds some further points. He is 'astonished that . . . scholars such as N.R. Hanson, Thomas Kuhn, Michael Polanyi, Paul Feyerabend, Sigmund Koch *et al.*, consider the distinction as invalid or at least misleading'.[3] And he points out that neither the psychology of invention nor any similarity, however great, between the sciences and the arts can show that it does not exist. In this he is certainly right. Even the most surprising stories about the manner in which scientists arrive at their theories cannot exclude the possibility that they proceed in an entirely different way once they have found them. *But this possibility is never realized.* Inventing theories and contemplating them in a relaxed and 'artistic' fashion, scientists often make moves that are forbidden by methodological rules. For example, they interpret the evidence so that it fits their fanciful ideas, eliminate difficulties by *ad hoc* procedures, push them aside, or simply refuse to take them seriously. The activities which according to Feigl belong to the context of discovery are, therefore, not just *different* from what philosophers say about justification, *they are in conflict with it.* Scientific practice does not contain two contexts moving *side by side*, it is a complicated *mixture* of

2. See the examples in Chapter 5.

3. 'Empiricism at Bay', MS, 1972, p. 2.

procedures, and we are faced by the question if this mixture should be left as it is, or if it should be replaced by a more 'orderly' arrangement. This is part one of the argument. Now we have seen that science as we know it today could not exist without a frequent overruling of the context of justification. This is part two of the argument. The conclusion is clear. Part one shows that we do not have a difference, but a mixture. Part two shows that replacing the mixture by an order that contains discovery on one side and justification on the other would have ruined science: we are dealing with a uniform practice all of whose ingredients are equally important for the growth of science. This disposes of the distinction.

A similar argument applies to the ritual distinction between methodological *prescriptions* and historical *descriptions*. Methodology, it is said, deals with what *should* be done and cannot be criticized by reference to *what is*. But we must of course make sure that our prescriptions have a *point of attack* in the historical material, and we must also make sure that their determined application leads to desirable results. We make sure by considering (historical, sociological, physical, psychological, etc.) *tendencies and laws* which tell us what is possible and what is not possible under the given circumstances and thus separate feasible prescriptions from those which are going to lead into dead ends. Again, progress can be made only if the distinction between the *ought* and the *is* is regarded as a temporary device rather than as a fundamental boundary line.

A distinction which once may have had a point but which has now definitely lost it is the distinction between *observational* terms and *theoretical* terms. It is now generally admitted that this distinction is not as sharp as it was thought to be only a few decades ago. It is also admitted, in complete agreement with Neurath's original views, that *both* theories *and* observations can be abandoned: theories may be removed because of conflicting observations, observations may be removed for theoretical reasons. Finally, we have discovered that *learning* does not go from observation to theory but always involves both elements. Experience arises *together* with theoretical assumptions not before them, and an experience without theory is just as incomprehensible as is (allegedly) a theory without experience: eliminate part of the theoretical knowledge of a sensing subject and you have a person who is completely disoriented and incapable of carrying out the simplest action. Eliminate further knowledge and his sensory world (his 'observation language') will start disintegrating, colours and other simple sensations will disappear, until

he is in a stage even more primitive than a small child. A small child, on the other hand, does not possess a stable perceptual world which he uses for making sense of the theories put before him. Quite the contrary – he passes through various perceptual stages which are only loosely connected with each other (earlier stages *disappear* when new stages take over – see Chapter 16) and which embody all the theoretical knowledge available at the time. Moreover, the whole process starts only because the child reacts correctly towards signals, *interprets them correctly*, because he possesses means of interpretation even before he has experienced his first clear sensation.

All these discoveries cry out for a new terminology that no longer separates what is so intimately connected in the development both of the individual and of science at large. Yet the distinction between observation and theory is still upheld. But what is its point? Nobody will deny that the sentences of science can be classified into long sentences and short sentences, or that its statements can be classified into those which are intuitively obvious and others which are not. Nobody will deny that such distinctions *can be made*. But nobody will put great weight on them, or will even mention them, *for they do not now play any decisive role in the business of science*. (This was not always so. Intuitive plausibility, for example, was once thought to be a most important guide to the truth; it disappeared from methodology the very moment intuition was replaced by experience, and by formal considerations.) Does experience play such a role? It does not, as we have seen. Yet the inference that the distinction between theory and observation has now ceased to be relevant, is either not drawn or is explicitly rejected.[4] Let us take a step forward and let us abandon this last trace of dogmatism in science!

Incommensurability, which I shall discuss next, is closely connected with the question of the rationality of science. Indeed one of the most general objections not merely to the *use of* incommensurable theories but even to the idea that *there are* such theories to be found in the history of science is the fear that they would severely restrict the efficacy of traditional, non-dialectical *argument*. Let us, therefore, look a little more closely at the critical *standards* which, according to some, constitute

4. 'Neurath fails to give . . . rules [which distinguish empirical statements from others] and thus unwittingly throws empiricism overboard', K.R. Popper, *The Logic of Scientific Discovery*, New York and London, 1959, p. 97.

the content of a 'rational' argument. More especially, let us look at the standards of the Popperian school, which are still being taken seriously in the more backward regions of knowledge. This will prepare us for the final step in our discussion of the issue between law-and-order methodologies and anarchism in science.

Some readers of my arguments in the above text have pointed out that Popper's 'critical' rationalism is sufficiently liberal to accommodate the developments I have described. Now critical rationalism is either a meaningful idea or it is a collection of slogans that can be adapted to any situation.

In the first case it must be possible to produce rules, standards, restrictions which permit us to separate critical behaviour (thinking, singing, writing of plays) from other types of behaviour so that we can *discover* irrational actions and *correct* them with the help of concrete suggestions. It is not difficult to produce the standards of rationality defended by the Popperian school.

These standards are standards of *criticism*: rational discussion consists in the attempt to criticize, and not in the attempt to prove or to make probable. Every step that protects a view from criticism, that makes it safe or 'well-founded', is a step away from rationality. Every step that makes it more vulnerable is welcome. In addition, it is recommended to abandon ideas which have been found wanting and it is forbidden to retain them in the face of strong and successful criticism unless one can present suitable counter-arguments. Develop your ideas so that they can be criticized; attack them relentlessly; do not try to protect them, but exhibit their weak spots; eliminate them as soon as such weak spots have become manifest – these are some of the rules put forth by our critical rationalists.

These rules become more definite and more detailed when we turn to the philosophy of science and, especially, to the philosophy of the natural sciences.

Within the natural sciences, criticism is connected with experiment and observations. The content of a theory consists in the sum total of those basic statements which contradict it; it is the class of its potential falsifiers. Increased content means increased vulnerability, hence theories of large content are to be preferred to theories of small content. Increase of content is welcome, decrease of content is to be avoided. A theory that contradicts an accepted basic statement must be given up. *Ad hoc*

hypotheses are forbidden – and so on. A science, however, that accepts the rules of a critical empiricism of this kind will develop in the following manner.

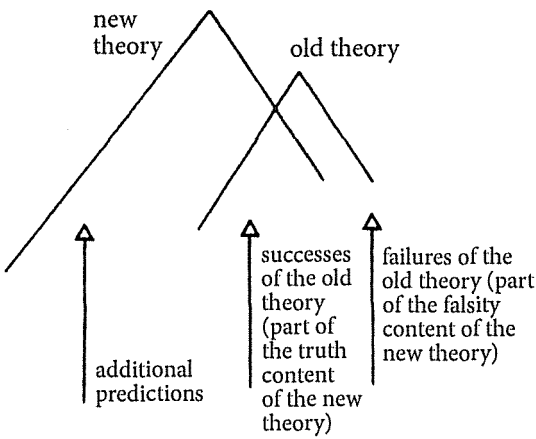
We start with a *problem*, such as the problem of the planets at the time of Plato. This problem (which I shall discuss in a somewhat idealized form) is not merely the result of *curiosity*, it is a *theoretical result*. It is due to the fact that certain *expectations* have been disappointed: on the one hand it seems to be clear that the stars must be divine, hence one expects them to behave in an orderly and lawful manner. On the other hand, one cannot find any easily discernible regularity. The planets, to all intents and purposes, move in a quite chaotic fashion. How can this fact be reconciled with the expectation and with the principles that underlie the expectation? Does it show that the expectation is mistaken? Or have we failed in our analysis of the facts? This is the problem.

It is important to see that the elements of the problem are not simply *given*. The 'fact' of irregularity, for example, is not accessible without further ado. It cannot be discovered by just anyone who has healthy eyes and a good mind. It is only through a certain expectation that it becomes an object of our attention. Or, to be more accurate, this fact of irregularity *exists* because there is an expectation of regularity and because there are ideas which define what it means to be 'regular'. After all, the term 'irregularity' makes sense only if we have a rule. In our case the rule that defines regularity asserts circular motion with constant angular velocity. The fixed stars agree with this rule and so does the sun, if we trace its path relative to the fixed stars. The planets do not obey the rule, neither directly, with respect to the earth, nor indirectly, with respect to the fixed stars.

(In the problem we are examining now the rule is formulated explicitly and it can be discussed. This is not always the case. Recognizing a colour as red is made possible by deep-lying patterns concerning the structure of our surroundings, and recognition does not occur when these patterns cease to exist.)

To sum up this part of the Popperian doctrine: research starts with a problem. The problem is the result of a conflict between an expectation and an observation which is constituted by the expectation. It is clear that this doctrine differs from the doctrine of inductivism where objective facts enter a passive mind and leave their traces there. It was prepared by Kant, Mach, Poincaré, Dingler, and by Mill (*On Liberty*).

Having formulated a problem, one tries to *solve* it. Solving a problem means inventing a theory that is relevant, falsifiable (to a degree larger than any alternative), but not yet falsified. In the case mentioned above (planets at the time of Plato), the problem is: to find circular motions of constant angular velocity for the purpose of saving the planetary phenomena. A first solution was provided by Eudoxos and then by Heracleides of Pontos.



Next comes the *criticism* of the theory that has been put forth in the attempt to solve the problem. Successful criticism removes the theory *once and for all* and creates a new problem, viz. to explain (a) why the theory was successful so far; (b) why it failed. Trying to solve *this* problem we need a new theory that reproduces the successful consequences of the older theory, denies its mistakes and makes additional predictions not made before. These are some of the *formal conditions* which a *suitable successor of a refuted theory* must satisfy. Adopting the conditions, one proceeds by conjecture and refutation from less general theories to more general theories and expands the content of human knowledge.

More and more facts are *discovered* (or constructed with the help of expectations) and are then explained by theories. There is no guarantee that scientists will solve every problem and replace every theory that has been refuted with a successor satisfying the formal conditions. The invention of theories depends on our talents and other fortuitous circumstances such as a satisfactory sex life. But as long as these talents

hold out, the enclosed scheme is a correct account of the growth of a knowledge that satisfies the rules of critical rationalism.

Now at this point, one may raise two questions.

1. Is it *desirable* to live in accordance with the rules of a critical rationalism?

2. Is it *possible* to have both a science as we know it and these rules?

As far as I am concerned, the first question is far more important than the second. True, science and related institutions play an important part in our culture, and they occupy the centre of interest for many philosophers (most philosophers are opportunists). Thus the ideas of the Popperian school were obtained by generalizing solutions for methodological and epistemological problems. Critical rationalism arose from the attempt to understand the Einsteinian revolution, and it was then extended to politics and even to the conduct of one's private life. Such a procedure may satisfy a *school philosopher*, who looks at life through the spectacles of his own technical problems and recognizes hatred, love, happiness, only to the extent to which they occur in these problems. But if we consider human interests and, above all, the question of human freedom (freedom from hunger, despair, from the tyranny of constipated systems of thought and *not* the academic 'freedom of the will'), then we are proceeding in the worst possible fashion.

For is it not possible that science as we know it today, or a 'search for the truth' in the style of traditional philosophy, will create a monster? Is it not possible that an objective approach that frowns upon personal connections between the entities examined will harm people, turn them into miserable, unfriendly, self-righteous mechanisms without charm and humour? 'Is it not possible,' asks Kierkegaard, 'that my activity as an objective [or a critico-rational] observer of nature will weaken my strength as a human being?'[5] I suspect the answer to many of these questions is affirmative and I believe that a reform of the sciences that makes them more anarchic and more subjective (in Kierkegaard's sense) is urgently needed.

5. *Papirer*, ed. Heiberg, VII, Pt. I, sec. A, No. 182. Mill tries to show how scientific method can be understood as part of a theory of man, and thus gives a positive answer to the question raised by Kierkegaard; see footnote 2 to Chapter 4.

But these are not the problems I want to discuss now. In the present essay I shall restrict myself to the second question and I shall ask: Is it possible to have both a science as we know it and the rules of a critical rationalism as just described? And to *this* question the answer seems to be a firm and resounding NO.

To start with we have seen, though rather briefly, that the actual development of institutions, ideas, practices, and so on, often *does not start from a problem* but rather from some extraneous activity, such as playing, which, as a side effect, leads to developments which later on can be interpreted as solutions to unrealized problems.[6] Are such developments to be excluded? And, if we do exclude them, will this not considerably reduce the number of our adaptive reactions and the quality of our learning process?

Secondly, we have seen, in Chapters 8–14, that a *strict principle of falsification*, or a 'naive falsificationism' as Lakatos calls it,[7] would wipe out science as we know it and would never have permitted it to start.

The demand for *increased content* is not satisfied either. Theories which effect the overthrow of a comprehensive and well-entrenched point of view, and take over after its demise, are initially restricted to a fairly narrow domain of facts, to a series of paradigmatic phenomena which lend them support, and they are only slowly extended to other areas. This can be seen from historical examples (footnote 12 of Chapter 8), and it is also plausible on general grounds: trying to develop a new theory, we must first take a *step back* from the evidence and reconsider the problem of observation (this was discussed in Chapter 11). Later on, of course, the theory is extended to other domains; but the mode of extension is only rarely determined by the elements that constitute the content of its predecessors. The slowly emerging conceptual apparatus of the theory *soon starts defining its own problems*, and earlier problems, facts, and observations are either forgotten or pushed aside as irrelevant. This is an entirely natural development, and quite unobjectionable. For

6. See the brief comments on the relation between idea and action in Chapter 1. For details see footnotes 31ff of 'Against Method', *Minnesota Studies*, Vol. 4, 1970.

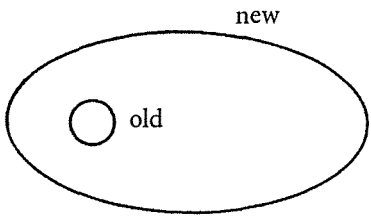
7. 'Falsification and the Methodology of Scientific Research Programmes', in Lakatos and Musgrave (eds), *Criticism and the Growth of Knowledge*, Cambridge, 1970, pp. 93ff. ('Naive falsificationism' is here also called 'dogmatic'.)

why should an ideology be constrained by older problems which, at any rate, make sense only in the abandoned context and which look silly and unnatural now? Why should it even *consider* the 'facts' that gave rise to problems of this kind or played a role in their solutions? Why should it not rather proceed in its own way, devising its own tasks and assembling its own domain of 'facts'? A comprehensive theory, after all, is supposed to contain also an *ontology* that determines what exists and thus delimits the domain of possible facts and possible questions. The development of science agrees with these considerations. New views soon strike out in new directions and frown upon the older *problems* (What is the base upon which the earth rests? What is the specific weight of phlogiston? What is the absolute velocity of the earth?) and the older *facts* (most of the facts described in the *Malleus Maleficarum* – Chapter 8, footnote 2 – the facts of Voodoo – Chapter 4, footnote 8 – the properties of phlogiston or those of the ether) which so much exercised the minds of earlier thinkers. And where they *do* pay attention to preceding theories, they try to accommodate their factual core in the manner already described, with the help of *ad hoc* hypotheses, *ad hoc* approximations, redefinition of terms, or by simply *asserting*, without any more detailed study of the matter, that the core 'follows from' the new basic principles.[8] They are 'grafted on to older programmes with which they [are] blatantly inconsistent'.[9]

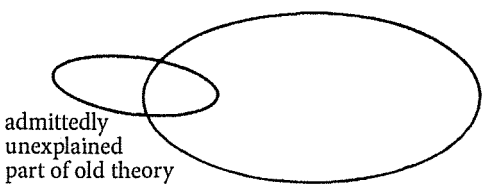
The result of all these procedures is an interesting *epistemological illusion*: the *imagined* content of the earlier theories (which is the intersection of the remembered consequences of these theories with the newly recognized domain of problems and facts) *shrinks* and may decrease to such an extent that it becomes smaller than the *imagined* content of the new ideologies (which are the actual consequences of these ideologies *plus* all those 'facts', laws, principles which are tied to them by *ad hoc* hypotheses, *ad hoc* approximations or by the say-so of some influential physicist or philosopher of science – and which properly belong to the predecessor). Comparing the old and the new it thus *appears* that the relation of empirical contents is like this

8. 'Einstein's theory is better than . . . Newton's theory *anno 1916* . . . *because* it explained everything that Newton's theory had successfully explained . . .', Lakatos, op. cit., p. 214.

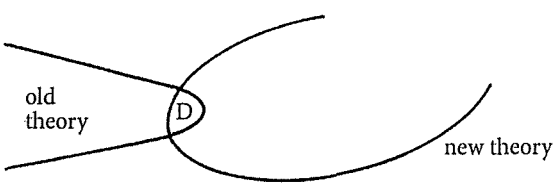
9. Lakatos, discussing Copernicus and Bohr, ibid., p. 143.



or, perhaps, like this



while in actual fact it is much more like this



domain D representing the problems and facts of the old theory which are still remembered and which have been distorted so as to fit into the new framework. It is this illusion which is responsible for the persistent survival of the demand for increased content.[10]

10. This illusion is the core of Elie Zahar's excellent paper on the development from Lorentz to Einstein. According to Zahar, Einstein superseded Lorentz with the explanation of the perihelion of Mercury (1915). But in 1915 nobody had as yet succeeded in giving a relativistic account of classical perturbation theory to the degree of approximation reached by Laplace and Poincaré, and the implications of Lorentz on the atomic level (electron theory of metals) were not accounted for either, but were gradually replaced by the quantum theory: Lorentz was 'superseded' not by one, but by at least two different and mutually incommensurable programmes. Lakatos, in his excellent reconstruction of the development of the research programme of Copernicus from the *Commentariolus* to the *De Revol.*, notes progressive changes but only because he omits the dynamical and the optical problems and concentrates on kinematics, pure and simple. Small wonder that both Zahar and Lakatos are under the impression that the content condition is still satisfied. See my short note 'Zahar on Einstein', in the *British Journal for the Philosophy of Science*, March

Finally, we have by now seen quite distinctly the need for *ad hoc hypotheses*: *ad hoc* hypotheses and *ad hoc* approximations create a tentative area of contact between 'facts' and those parts of a new view which seem capable of explaining them, at some time in the future and after addition of much further material. They specify possible explananda and explanatia, and thus determine the direction of future research. They may have to be retained forever if the new framework is partly unfinished (this happened in the case of the quantum theory, which needs the classical concepts to turn it into a complete theory). Or they are incorporated into the new theory as theorems, leading to a redefinition of the basic terms of the proceeding ideology (this happened in the cases of Galileo and of the theory of relativity). The demand that the truth-content of the earlier theory as *conceived while the earlier theory reigned supreme* be included in the truth-content of the successor is violated in either case.

To sum up: wherever we look, whatever examples we consider, we see that the principles of critical rationalism (take falsifications seriously; increase content; avoid *ad hoc* hypotheses; 'be honest' – whatever *that* means; and so on) and, *a fortiori*, the principles of logical empiricism (be precise; base your theories on measurements; avoid vague and untestable ideas; and so on), though practised in special areas, give an inadequate account of the past development of science as a whole and are liable to hinder it in the future. They give an inadequate account of science because science is much more 'sloppy' and 'irrational' than its methodological image. And they are liable to hinder it because the attempt to make science more 'rational' and more precise is bound to wipe it out, as we have seen. The difference between science and methodology which is such an obvious fact of history, therefore, indicates a weakness of the latter, and perhaps of the 'laws of reason' as well. For what appears as 'sloppiness', 'chaos' or 'opportunism' when compared with such laws has a most important function in the development of those very theories which we today regard as essential parts of our knowledge of nature. *These 'deviations', these 'errors', are preconditions of progress.* They permit knowledge to survive in the complex and difficult world which we inhabit, they permit *us* to remain free and happy agents. Without 'chaos', no knowledge. Without a frequent dismissal of reason, no progress. Ideas

1974 as well as R.N. Nugaev, 'Special Relativity as a Stage in the Development of Quantum Theory', *Historia Scientarium*, No. 34, 1988, pp. 57ff.

which today form the very basis of science exist only because there were such things as prejudice, conceit, passion; because these things *opposed reason*; and because they *were permitted to have their way*. We have to conclude, then, that *even within* science reason cannot and should not be allowed to be comprehensive and that it must often be overruled, or eliminated, in favour of other agencies. There is not a single rule that remains valid under all circumstances and not a single agency to which appeal can always be made.[11]

11. Even Lakatos' ingenious methodology does not escape this indictment. Lakatos seems liberal because he forbids very little and he seems rational because he still forbids something. But the only thing he forbids is to *describe* a 'degenerating research programme', i.e. a research programme lacking in novel predictions and cluttered with *ad hoc* adaptations, as progressive. He does not forbid its use. But this means that his standards permit a criminal to commit as many crimes as he wants provided he never lies about them. Details in my *Philosophical Papers*, Vol. 2, Chapter 10.